

CRITICAL POINTS OF RANDOM POLYNOMIALS AND CHARACTERISTIC POLYNOMIALS OF RANDOM MATRICES

SEAN O'ROURKE

ABSTRACT. Let p_n be the characteristic polynomial of an $n \times n$ random matrix drawn from one of the compact classical matrix groups. We show that the critical points of p_n converge to the uniform distribution on the unit circle as n tends to infinity. More generally, we show the same limit for a class of random polynomials whose roots lie on the unit circle. Our results extend the work of Pemantle–Rivin [30] and Kabluchko [19] to the setting where the roots are neither independent nor identically distributed.

1. INTRODUCTION

A *critical point* of a polynomial f is a root of its derivative f' . There are many results concerning the location of critical points of polynomials whose roots are known. For example, the famous Gauss–Lucas theorem offers a geometric connection between the roots of a polynomial and the roots of its derivative.

Theorem 1 (Gauss–Lucas; Theorem 6.1 from [24]). *If f is a non-constant polynomial with complex coefficients, then all zeros of f' belong to the convex hull of the set of zeros of f .*

There are also a number of refinements of Theorem 1. We refer the reader to [1, 3, 5, 8, 10, 14, 18, 22, 23, 25, 29, 32, 34, 35, 39, 42] and references therein.

Pemantle and Rivin [30] initiated the study of a probabilistic version of the Gauss–Lucas theorem. In order to introduce their results, we fix the following notation. For a polynomial f of degree n , we define the empirical measure constructed from the roots of f as

$$\mu_f := \frac{1}{n} \sum_{z \in \mathbb{C}: f(z)=0} \mathcal{N}_f(z) \delta_z,$$

where $\mathcal{N}_f(z)$ is the multiplicity of the zero at z and δ_z is the unit point mass at z .

For an integer $k \geq 1$, we use the convention that

$$\mu_f^{(k)} := \mu_{f^{(k)}}.$$

That is, $\mu_f^{(k)}$ is the empirical measure constructed from the roots of the k -th derivative of f . Similarly, we write μ'_f to denote the empirical measure constructed from the critical points of f .

Let X_1, X_2, \dots be independent and identically distributed (iid) random variables taking values in \mathbb{C} . Let μ be the probability distribution of X_1 . For each $n \geq 1$, consider the polynomial

$$p_n(z) := (z - X_1) \cdots (z - X_n). \tag{1}$$

Pemantle and Rivin [30] show, assuming μ has finite one-dimensional energy, that μ'_{p_n} converges weakly to μ as n tends to infinity.

Let us recall what it means for a sequence of random probability measures to converge weakly.

Definition 2 (Weak convergence of random probability measures). Let T be a topological space (such as \mathbb{R} or \mathbb{C}), and let \mathcal{B} be its Borel σ -field. Let $(\mu_n)_{n \geq 1}$ be a sequence of random probability measures on (T, \mathcal{B}) , and let μ be a probability measure on (T, \mathcal{B}) . We say μ_n *converges weakly to μ in probability* as $n \rightarrow \infty$ (and write $\mu_n \rightarrow \mu$ in probability) if for all bounded continuous $f : T \rightarrow \mathbb{R}$ and any $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \int f d\mu_n - \int f d\mu \right| > \varepsilon \right) = 0.$$

In other words, $\mu_n \rightarrow \mu$ in probability as $n \rightarrow \infty$ if and only if $\int f d\mu_n \rightarrow \int f d\mu$ in probability for all bounded continuous $f : T \rightarrow \mathbb{R}$. Similarly, we say μ_n *converges weakly to μ almost surely* as $n \rightarrow \infty$ (and write $\mu_n \rightarrow \mu$ almost surely) if for all bounded continuous $f : T \rightarrow \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu$$

almost surely.

Kabluchko [19] generalized the results of Pemantle and Rivin to the following.

Theorem 3 (Kabluchko). *Let μ be any probability measure on \mathbb{C} . Let X_1, X_2, \dots be a sequence of iid random variables with distribution μ . For each $n \geq 1$, let p_n be the degree n polynomial given in (1). Then μ'_{p_n} converges weakly to μ in probability as $n \rightarrow \infty$.*

The following corollary of Theorem 3 will be relevant to this note.

Corollary 4. *Let $\theta_1, \theta_2, \dots$ be a sequence of iid random variables distributed uniformly on $[0, 2\pi)$. For each $n \geq 1$, let*

$$p_n(z) := \prod_{j=1}^n (z - e^{i\theta_j}).$$

Then μ'_{p_n} converges in probability to the uniform probability distribution on the unit circle centered at the origin in the complex plane as $n \rightarrow \infty$.

Corollary 4 also follows from the work of Subramanian in [40].

2. MAIN RESULTS

The goal of this note is to prove a version of Theorem 3 when the random variables X_1, X_2, \dots are neither independent nor identically distributed. Of particular interest will be the case when the roots of p_n are eigenvalues of a random matrix.

The *eigenvalues* of a square matrix M are the zeros of its characteristic polynomial $p_M(z) := \det(zI - M)$, where I denotes the identity matrix. We let μ_M denote the empirical spectral measure of M . That is, μ_M is the empirical measure constructed from the roots of the characteristic polynomial p_M . Similarly, we let μ'_M be the empirical measure constructed from the roots of p'_M .

As a motivating example, we begin with the case when M is Hermitian.

2.1. Characteristic polynomials of Hermitian random matrices. If the matrix M is Hermitian (that is, $M = M^*$, where M^* denotes the conjugate transpose of M), then the eigenvalues of M are real and μ_M is a probability measure on the real line. In this case, Theorem 6 below describes the well-known connection between μ_M and μ'_M . Before stating the result, we first recall the following definition.

Definition 5 (Lévy distance). Let μ and ν be two probability measures on the real line with cumulative distribution functions F and G respectively. Then the *Lévy distance* $L(\mu, \nu)$ between μ and ν is given by

$$L(\mu, \nu) := \inf\{\varepsilon \geq 0 : G(x - \varepsilon) - \varepsilon \leq F(x) \leq G(x + \varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\}.$$

It is well-known, for measures on the real line, that convergence in Lévy distance is equivalent to convergence in distribution; we refer the reader to [21, Chapter 13.2] and [21, Exercise 13.2.6] for further details.

Theorem 6. *For each $n \geq 1$, let X_n be a $n \times n$ random Hermitian matrix. Then μ'_{X_n} is a random probability measure on the real line and*

$$L(\mu_{X_n}, \mu'_{X_n}) \longrightarrow 0$$

almost surely as $n \rightarrow \infty$

Proof. Since the eigenvalues of X_n are real, the Gauss–Lucas theorem (Theorem 1) guarantees that μ'_{X_n} is a probability measure on the real line.

Let $I \subset \mathbb{R}$ be an interval. Let N_I denote the number of zeros of p_{X_n} in I (i.e. the number of eigenvalues of X_n in I), and let N'_I denote the number of critical points of p_{X_n} in I . Since the zeros of p'_{X_n} interlace the zeros of p_{X_n} , we have

$$|N_I - N'_I| \leq 1,$$

and the claim follows from [2, Lemma B.18]. \square

As a concrete example, we present the following corollary for Wigner random matrices.

Corollary 7. *Let ξ be a complex-valued random variable with unit variance, and let ζ be a real-valued random variable. For each $n \geq 1$, let X_n be a $n \times n$ Hermitian matrix whose diagonal entries are iid copies of ζ , those above the diagonal are iid copies of ξ , and all the entries on and above the diagonal are independent. Then $\mu'_{\frac{1}{\sqrt{n}}X_n}$ is a probability measure on the real line and*

$$\mu'_{\frac{1}{\sqrt{n}}X_n} \longrightarrow \mu_{\text{sc}}$$

almost surely as $n \rightarrow \infty$, where μ_{sc} is the measure on the real line with density

$$\rho_{\text{sc}}(x) := \begin{cases} \frac{1}{2\pi} \sqrt{4 - x^2}, & |x| \leq 2 \\ 0, & |x| > 2. \end{cases}$$

Proof. In view of [2, Theorem 2.5] and [21, Exercise 13.2.6], we have

$$L\left(\mu_{\frac{1}{\sqrt{n}}X_n}, \mu_{\text{sc}}\right) \longrightarrow 0$$

almost surely as $n \rightarrow \infty$. Thus, the claim follows from Theorem 6 by applying the triangle inequality for Lévy distance. \square

The same arguments can also be used to generalize Theorem 6 and Corollary 7 to higher-order derivatives.

2.2. Random matrices from the compact classical groups. In this note, we extend Theorem 6 to random matrices which are not Hermitian. In particular, we consider random matrices distributed according to Haar measure on the compact classical matrix groups. We begin by recalling some definitions.

Definition 8 (Compact classical matrix groups).

- (1) An $n \times n$ matrix M over \mathbb{R} is *orthogonal* if

$$MM^T = M^T M = I_n,$$

where I_n denotes the $n \times n$ identity matrix and M^T is the transpose of M . The set of $n \times n$ orthogonal matrices over \mathbb{R} is denoted by $O(n)$.

- (2) The set $SO(n) \subset O(n)$ of *special orthogonal matrices* is defined by

$$SO(n) := \{M \in O(n) : \det(M) = 1\}.$$

- (3) An $n \times n$ matrix M over \mathbb{C} is *unitary* if

$$MM^* = M^* M = I_n,$$

where M^* denotes the conjugate transpose of M . The set of $n \times n$ unitary matrices over \mathbb{C} is denoted $U(n)$.

- (4) If n is even, we say an $n \times n$ matrix M over \mathbb{C} is *symplectic* if $M \in U(n)$ and

$$MJM^* = M^* JM = J,$$

where

$$J := \begin{bmatrix} 0 & I_{n/2} \\ -I_{n/2} & 0 \end{bmatrix}.$$

The set of $n \times n$ symplectic matrices over \mathbb{C} is denoted $Sp(n)$.

Recall that if M is a matrix from one of the compact matrix groups introduced above, then the eigenvalues of M all lie on the unit circle in the complex plane centered at the origin.

For any compact Lie group G , there exists a unique translation-invariant probability measure on G called *Haar measure*; see, for example, [12, Chapter 2.2]. In this note, we will be interested in the case when G is one of classical compact matrix groups defined above.

For the compact matrix groups, there are a number of intuitive ways to describe a matrix distributed according to Haar measure. Recall that a complex standard normal random variable Z can be represented as $Z = X + iY$, where X and Y are independent real normal random variables with mean zero and variance $1/2$. Form an $n \times n$ random matrix with independent complex standard normal entries and perform the Gram–Schmidt algorithm on the columns. The result is a random unitary matrix distributed according to Haar measure on $U(n)$. Indeed, invariance follows from the invariance of complex Gaussian random vectors under $U(n)$. Similar Gaussian constructions yield random matrices distributed according to Haar measure on the other compact matrix groups.

We now present our main result for the classical compact matrix groups.

Theorem 9. *For each $n \geq 1$, let M_n be an $n \times n$ matrix Haar distributed on $O(n)$, $SO(n)$, $U(n)$, or $Sp(n)$. Then μ'_{M_n} converges in probability as $n \rightarrow \infty$ to the uniform probability distribution on the unit circle centered at the origin.*

Remark 10. If M_n is an $n \times n$ random matrix Haar distributed on $O(n)$, $SO(n)$, $U(n)$, or $Sp(n)$, then μ_{M_n} also converges in probability to the uniform distribution on the unit circle centered at the origin. Moreover, in [27], the authors prove that the convergence holds in the almost sure sense and give a rate of convergence.

Figure 1 depicts a numerical simulation of the zeros and critical points of the characteristic polynomial of a random orthogonal matrix chosen according to Haar measure.

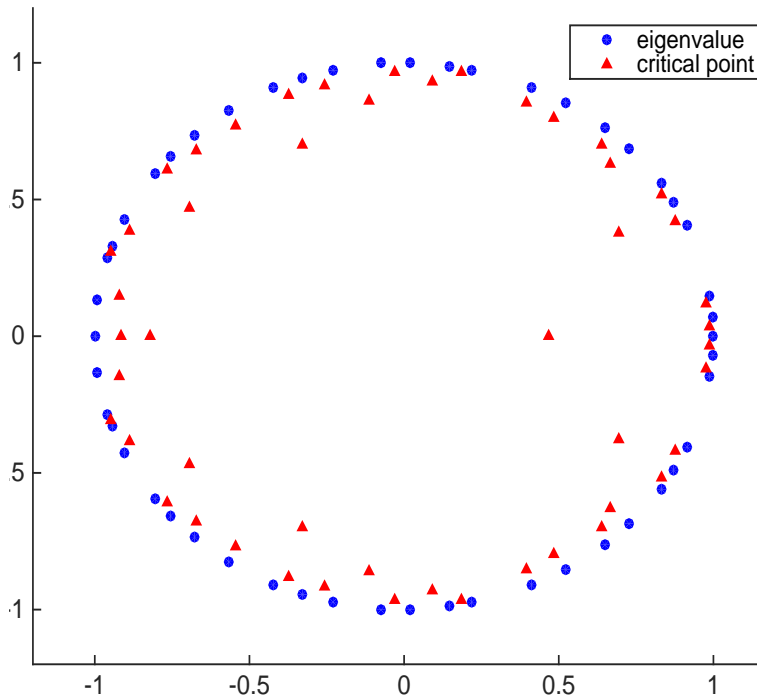


FIGURE 1. The zeros and critical points of the characteristic polynomial of a random orthogonal matrix of size 50×50 .

2.3. Random polynomials with roots on the unit circle. More generally, we consider random polynomials of the form

$$p_n(z) := \prod_{j=1}^n (z - X_j),$$

where X_1, X_2, \dots are random variables on the unit circle, not necessarily independent or identically distributed. Indeed, we will deduce Theorem 9 from the following more general result.

Theorem 11. *For each $n \geq 1$, let $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ be random variables on $[0, 2\pi)$. Set*

$$p_n(z) := \prod_{j=1}^n (z - e^{i\theta_j^{(n)}}).$$

Assume

(i) *we have*

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) = 0,$$

(ii) *for almost every $z \in \mathbb{D} := \{w \in \mathbb{C} : |w| < 1\}$,*

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{m=0}^{\lfloor \log^2 n \rfloor} z^m \sum_{j=1}^n e^{-i\theta_j^{(n)}(m+1)} \right| \leq \delta \right) = 0,$$

(iii) *for all integers $m \geq 1$,*

$$\frac{1}{n} \sum_{j=1}^n e^{i\theta_j^{(n)}m} \longrightarrow 0$$

in probability as $n \rightarrow \infty$.

Then μ'_{p_n} converges in probability as $n \rightarrow \infty$ to the uniform probability distribution on the unit circle centered at the origin.

We pause for a moment to discuss the three assumptions of Theorem 11. Roughly speaking, condition (iii) is the most important, while conditions (i) and (ii) are technical anti-concentration estimates. Indeed, condition (iii) implies that the empirical measure constructed from $e^{i\theta_1^{(n)}}, \dots, e^{i\theta_n^{(n)}}$ converges in probability as $n \rightarrow \infty$ to the uniform distribution on the unit circle centered at the origin. We also note that the sequence $\log^2 n$ appearing in condition (ii) is not vital; it can be replaced with $(\log n)^{1+\varepsilon}$ for any $\varepsilon > 0$.

We will also verify the following alternative formulation of Theorem 11.

Theorem 12 (Alternative formulation). *For each $n \geq 1$, let $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ be random variables on $[0, 2\pi)$. Set*

$$p_n(z) := \prod_{j=1}^n (z - e^{i\theta_j^{(n)}}).$$

Assume

(i) *we have*

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) = 0,$$

(ii) *for almost every $z \in \mathbb{D} := \{w \in \mathbb{C} : |w| < 1\}$,*

$$\frac{1}{n} \log \left| \sum_{j=1}^n \frac{1}{z - e^{i\theta_j^{(n)}}} \right| \longrightarrow 0$$

in probability as $n \rightarrow \infty$,

(iii) for all integers $m \geq 1$,

$$\frac{1}{n} \sum_{j=1}^n e^{i\theta_j^{(n)} m} \longrightarrow 0$$

in probability as $n \rightarrow \infty$.

Then μ'_{p_n} converges in probability as $n \rightarrow \infty$ to the uniform probability distribution on the unit circle centered at the origin.

We will use Theorem 11 to prove Theorem 9. However, Theorem 12 is also useful. For example, we can recover Corollary 4 from Theorem 12. Indeed, if $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ are iid random variables uniformly distributed on $[0, 2\pi)$, then the assumptions of Theorem 12 can be verified using [31, Theorem 2.22], [19, Lemma 2.1], and the law of large numbers. Theorem 12 is also useful when the random variables $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ are dependent. To illustrate this point, we will use Theorem 12 to verify the following corollary.

Corollary 13. *Let $\theta_1, \theta_2, \dots$ be a sequence of iid random variables distributed uniformly on $[0, 2\pi)$. For each $n \geq 1$, set*

$$p_{2n}(z) := \prod_{j=1}^n (z - e^{i\theta_j})(z - e^{-i\theta_j}).$$

Then $\mu'_{p_{2n}}$ converges in probability as $n \rightarrow \infty$ to the uniform probability distribution on the unit circle centered at the origin.

2.4. Discussion and open problems. We conjecture that for many classes of random polynomials the critical points should be stochastically close to the distribution of the roots. Intuitively, this would imply that the distribution of the critical points is nearly identical to the distribution of the roots for a “typical” polynomial of high degree.

As another example, consider the Kac polynomials. In this case, one can show even more by applying the results of Kabluchko and Zaporozhets [20].

Theorem 14 (Kabluchko–Zaporozhets). *Let ξ_0, ξ_1, \dots be a sequence of non-degenerate iid random variables such that $\mathbb{E} \log(1 + |\xi_0|) < \infty$. For each $n \geq 1$, let $f_n(z) = \sum_{j=0}^n \xi_j z^j$. Fix an integer $k \geq 1$. Then μ_{f_n} and $\mu_{f_n}^{(k)}$ both converge in probability as $n \rightarrow \infty$ to the uniform probability distribution on the unit circle centered at the origin.*

Proof. Both claims follow from [20, Theorem 2.2] by simply estimating the coefficients of $f_n^{(k)}$. In fact, a similar argument allows one to consider solutions of the equation $f_n^{(k)} = c_n$, where c_n is a constant; see [20, Remark 2.11] for details. \square

We conjecture that this universality phenomenon should also hold for the characteristic polynomial of many random matrix ensembles. For instance, Figure 2 depicts a numerical simulation of the zeros and critical points of the characteristic polynomial of a random matrix with iid real standard normal entries.

2.5. Organization. The paper is organized as follows. In Section 3, we prove Theorem 9 and Corollary 13 using Theorems 11 and 12. The proof of Theorems 11 and 12 is contained in Sections 4 and 5.

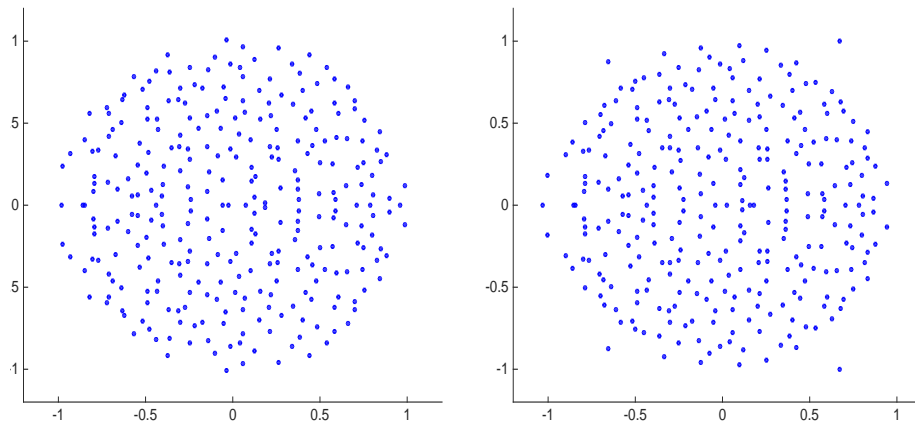


FIGURE 2. The roots and critical points of the characteristic polynomial of an $n \times n$ random matrix with iid real standard normal entries when $n = 300$. The figure on the left depicts the location of the eigenvalues (scaled by $1/\sqrt{n}$). The figure on the right shows the location of the critical points (scaled by $1/\sqrt{n}$).

2.6. Notation. We let $\mathbb{D}_r := \{z \in \mathbb{C} : |z| < r\}$ be the open disk of radius $r > 0$ centered at the origin and $\overline{\mathbb{D}}_r := \{z \in \mathbb{C} : |z| \leq r\}$ its closure. We write $\mathbb{D} := \mathbb{D}_1$.

We let C and K denote constants that are non-random and may take on different values from one appearance to the next. The notation K_p means that the constant K depends on another parameter p .

We write a.s., a.a., and a.e. for almost surely, Lebesgue almost all, and Lebesgue almost everywhere respectively. For an event E , we let $\mathbf{1}_E$ denote the indicator function of E ; E^C is the complement of E .

3. PROOF OF THEOREM 9 AND COROLLARY 13

In this section, we prove Theorem 9 and Corollary 13 using Theorems 11 and 12.

3.1. Proof of Theorem 9. We will apply Theorem 11 to prove Theorem 9. For each $n \geq 1$, let M_n be an $n \times n$ matrix Haar distributed on $O(n)$, $SO(n)$, $U(n)$, or $Sp(n)$. Let $e^{i\theta_1^{(n)}}, \dots, e^{i\theta_n^{(n)}}$ be the eigenvalues of M_n , where $\theta_1^{(n)}, \dots, \theta_n^{(n)} \in [0, 2\pi)$. It now suffices to show that the eigenvalues of M_n satisfy the three assumptions of Theorem 11.

In order to verify the assumptions of Theorem 11, we will need the following multivariate central limit theorem for traces of random matrices from the classical matrix groups found in [9, 38]. First, we recall the Wasserstein distance between two probability distributions.

Definition 15 (Wasserstein distance). Let (S, d) be a separable metric space, and let μ and ν be two probability measures on S . By $M(\mu, \nu)$ we denote the set of all probability measures on $S \times S$ with marginals μ and ν . The *Wasserstein distance*

$d_{\mathcal{W}}(\mu, \nu)$ between μ and ν is defined by

$$d_{\mathcal{W}}(\mu, \nu) := \inf \left\{ \int d(x, y) d\pi(x, y) : \pi \in M(\mu, \nu) \right\}.$$

We write $d_{\mathcal{W}}(P, Q)$, where P and Q are two random variables taking values in S , to mean the Wasserstein distance between the distributions of P and Q .

The Kantorovich–Rubinstein theorem gives an equivalent formulation of the Wasserstein distance in terms of Lipschitz functions on the separable metric space (S, d) . We refer the reader to [11, Section 11.8] for further details. We now state the results from [9, 38]; the case where M_n is drawn according to Haar measure from $U(n)$, $SO(n)$, or $Sp(n)$ is handled in [9, Theorem 1.1], while the orthogonal group $O(n)$ is studied in [38, Theorem 5.1].

Theorem 16 (Döbler–Stolz). *Let M_n be distributed according to Haar measure on $O(n)$, $SO(n)$, $U(n)$, or $Sp(n)$. For integers $d \geq 1$ and $r = 1, \dots, d$, consider the r -dimensional (complex or real) random vector*

$$W_{d,r,n} := (f_{d-r+1}(M_n), f_{d-r+2}(M_n), \dots, f_d(M_n))^T,$$

where $f_j(M_n) := \text{tr}(M_n^j)$ in the unitary case,

$$f_j(M_n) := \begin{cases} \text{tr}(M_n^j), & \text{if } j \text{ is odd,} \\ \text{tr}(M_n^j) - 1, & \text{if } j \text{ is even} \end{cases}$$

in the orthogonal and special orthogonal cases, and

$$f_j(M_n) := \begin{cases} \text{tr}(M_n^j), & \text{if } j \text{ is odd,} \\ \text{tr}(M_n^j) + 1, & \text{if } j \text{ is even} \end{cases}$$

in the symplectic case. In the orthogonal, special orthogonal, and symplectic cases, let $Z_{r,d} := (Z_{d-r+1}, \dots, Z_d)^T$ denote an r -dimensional real standard normal random vector. In the unitary case, Z is defined as a standard complex normal random vector. In all cases take Σ to be the diagonal matrix $\text{diag}(d-r+1, d-r+2, \dots, d)$, and write $Z_{\Sigma,r,d} := \Sigma^{1/2} Z_{r,d}$. Then there exists an absolute constant $C > 0$ (independent of r , d , and n) such that, for any $n \geq 4d+1$, we have

$$d_{\mathcal{W}}(W_{n,r,d}, Z_{\Sigma,r,d}) \leq C \frac{\max \left\{ \frac{r^{7/2}}{(d-r+1)^{3/2}}, (d-r)^{3/2} \sqrt{r} \right\}}{n}.$$

Remark 17. There is a large collection of literature concerning traces of random elements from the classical compact matrix groups. We refer the reader to [6, 7, 9, 13, 16, 17, 28, 36, 37, 38] and references therein.

We now verify the three assumptions of Theorem 11. We will use the same notation as in Theorem 16. Set $N := \lfloor \log^2 n \rfloor$. By Theorem 16, there exists random variables $\xi_1^{(n)}, \dots, \xi_{N+1}^{(n)}$ such that, for n sufficiently large, the random vector

$$(\xi_1^{(n)}, \dots, \xi_{N+1}^{(n)})^T$$

has the same distribution as

$$(f_1(M_n), \dots, f_{N+1}(M_n))^T,$$

and

$$\mathbb{E} \sqrt{\sum_{m=0}^N \left| \xi_{m+1}^{(n)} - \sqrt{m+1} Z_{m+1} \right|^2} \leq C \frac{(N+1)^{7/2}}{n}, \quad (2)$$

where $Z := (Z_1, \dots, Z_{N+1})^T$ is a standard normal random vector. Here $f_j(M_n)$ is defined as in Theorem 16, $C > 0$ is an absolute constant, and Z is a complex standard normal random vector in the unitary case and a real standard normal random vector in the other cases.

For any positive integer m , we write

$$\operatorname{tr} M_n^m = f_m(M_n) + \alpha_{n,m},$$

where $\alpha_{n,m}$ is deterministic and can take the values ± 1 or 0 depending on whether m is even or odd and depending on which classical matrix group M_n is drawn from.

We now verify condition (iii) of Theorem 11. Let m be a positive integer. For any $\eta > 0$ and all n sufficiently large, by Markov's inequality, we have

$$\begin{aligned} \mathbb{P}\left(\frac{1}{n} |\operatorname{tr} M_n^m| > \eta\right) &= \mathbb{P}(|f_m(M_n) + \alpha_{n,m}| > n\eta) \\ &= \mathbb{P}\left(|\xi_m^{(n)} + \alpha_{n,m}| > n\eta\right) \\ &\leq \mathbb{P}\left(|\xi_m^{(n)}| > \frac{n\eta}{2}\right) \\ &\leq 2 \frac{\mathbb{E}|\xi_m^{(n)}|}{n\eta} \\ &\leq 2 \frac{\mathbb{E}|\xi_m^{(n)} - \sqrt{m}Z_m|}{n\eta} + 2\sqrt{m} \frac{\mathbb{E}|Z_m|}{n\eta}. \end{aligned}$$

Therefore, by (2), we conclude that, for any $\eta > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{n} |\operatorname{tr} M_n^m| > \eta\right) = 0,$$

which completes the verification of condition (iii). (Alternatively, condition (iii) also follows from the results in [27].)

It remains to verify conditions (i) and (ii) of Theorem 11. Notice that condition (i) follows from condition (ii) in the case that $z = 0$. Thus, it suffices to prove condition (ii) for all $z \in \mathbb{D}$.

To this end, define the event

$$E_n := \left\{ \sqrt{\sum_{m=0}^N \left| \xi_{m+1}^{(n)} - \sqrt{m+1} Z_{m+1} \right|^2} \leq \frac{1}{100 \log^2 n} \right\}.$$

By Markov's inequality and (2), it follows that

$$\lim_{n \rightarrow \infty} \mathbb{P}(E_n^C) = 0. \tag{3}$$

It remains to show that, for all $z \in \mathbb{D}$,

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\left| \sum_{m=0}^N z^m \operatorname{tr} M_n^{m+1} \right| \leq \delta\right) = 0.$$

By symmetry, it suffices to show that for all $z \in \mathbb{D}$,

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\left| \sum_{m=0}^N z^m \operatorname{tr} M_n^{m+1} \right| \leq \delta\right) = 0.$$

Fix $z \in \mathbb{D}$. Observe that

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{m=0}^N z^m \operatorname{tr} M_n^{m+1} \right| \leq \delta \right) &= \mathbb{P} \left(\left| \sum_{m=0}^N z^m (\xi_{m+1}^{(n)} + \alpha_{n,m+1}) \right| \leq \delta \right) \\ &\leq \mathbb{P} \left(\left\{ \left| \sum_{m=0}^N z^m (\xi_{m+1}^{(n)} + \alpha_{n,m+1}) \right| \leq \delta \right\} \cap E_n \right) + \mathbb{P}(E_n^C). \end{aligned}$$

Thus, by (3), it suffices to show that

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left\{ \left| \sum_{m=0}^N z^m (\xi_{m+1}^{(n)} + \alpha_{n,m+1}) \right| \leq \delta \right\} \cap E_n \right) = 0. \quad (4)$$

Notice that on the event E_n , we have

$$\begin{aligned} &\left| \sum_{m=0}^N z^m (\xi_{m+1}^{(n)} + \alpha_{n,m+1}) - \sum_{m=0}^N z^m (\sqrt{m+1} Z_{m+1} + \alpha_{n,m+1}) \right| \\ &\leq \sum_{m=0}^N \left| \xi_{m+1}^{(n)} - \sqrt{m+1} Z_{m+1} \right| \\ &\leq \sqrt{N+1} \sqrt{\sum_{m=0}^N \left| \xi_{m+1}^{(n)} - \sqrt{m+1} Z_{m+1} \right|^2} \\ &\leq \frac{1}{\log n} \end{aligned}$$

by the Cauchy–Schwarz inequality.

Thus, by considering just the real part, we conclude that

$$\begin{aligned} &\mathbb{P} \left(\left\{ \left| \sum_{m=0}^N z^m (\xi_{m+1}^{(n)} + \alpha_{n,m+1}) \right| \leq \delta \right\} \cap E_n \right) \\ &\leq \mathbb{P} \left(\left| \sum_{m=0}^N z^m (\sqrt{m+1} Z_{m+1} + \alpha_{n,m+1}) \right| \leq \delta + \frac{1}{\log n} \right) \\ &\leq \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} (\operatorname{Re}(z^m) \operatorname{Re}(Z_{m+1}) - \operatorname{Im}(z^m) \operatorname{Im}(Z_{m+1})) + \operatorname{Re}(z^m) \alpha_{n,m+1} \right| \leq \delta + \frac{1}{\log n} \right) \\ &\leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} (\operatorname{Re}(z^m) \operatorname{Re}(Z_{m+1}) - \operatorname{Im}(z^m) \operatorname{Im}(Z_{m+1})) + x \right| \leq \delta + \frac{1}{\log n} \right). \end{aligned}$$

We now consider two cases. In the orthogonal, special orthogonal, or symplectic cases, we observe that $\operatorname{Im}(Z_{m+1}) = 0$ for $m = 0, \dots, N$. In this case, we have

$$\begin{aligned} &\sup_{x \in \mathbb{R}} \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} (\operatorname{Re}(z^m) \operatorname{Re}(Z_{m+1}) - \operatorname{Im}(z^m) \operatorname{Im}(Z_{m+1})) + x \right| \leq \delta + \frac{1}{\log n} \right) \\ &= \sup_{x \in \mathbb{R}} \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} \operatorname{Re}(z^m) Z_{m+1} + x \right| \leq \delta + \frac{1}{\log n} \right) \\ &= \sup_{x \in \mathbb{R}} \mathbb{P} \left(|\sigma_z Z_1 + x| \leq \delta + \frac{1}{\log n} \right), \end{aligned}$$

where

$$\sigma_z^2 := \sum_{m=0}^N |\operatorname{Re}(z^m)|^2 (m+1) \geq 1.$$

Here we used that Z_1, \dots, Z_{N+1} are iid real standard normal random variables, and hence any linear combination of Z_1, \dots, Z_{N+1} is also normal. Thus, we conclude that

$$\begin{aligned} \sup_{x \in \mathbb{R}} \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} (\operatorname{Re}(z^m) \operatorname{Re}(Z_{m+1}) - \operatorname{Im}(z^m) \operatorname{Im}(Z_{m+1})) + x \right| \leq \delta + \frac{1}{\log n} \right) \\ \leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(|Z_1 + x| \leq \delta + \frac{1}{\log n} \right) \\ \leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(|Z_1 + x| \leq \sqrt{2}\delta + \frac{\sqrt{2}}{\log n} \right). \end{aligned}$$

For the unitary case, we observe that

$$\begin{aligned} \sup_{x \in \mathbb{R}} \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} (\operatorname{Re}(z^m) \operatorname{Re}(Z_{m+1}) - \operatorname{Im}(z^m) \operatorname{Im}(Z_{m+1})) + x \right| \leq \delta + \frac{1}{\log n} \right) \\ \leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(|\sigma'_z \operatorname{Re}(Z_1) + x| \leq \delta + \frac{1}{\log n} \right), \end{aligned}$$

where

$$\sigma_z'^2 := \sum_{m=0}^N |z|^{2m} (m+1) \geq 1.$$

Thus, by the same reasoning as in the other cases, we conclude that

$$\begin{aligned} \sup_{x \in \mathbb{R}} \mathbb{P} \left(\left| \sum_{m=0}^N \sqrt{m+1} (\operatorname{Re}(z^m) \operatorname{Re}(Z_{m+1}) - \operatorname{Im}(z^m) \operatorname{Im}(Z_{m+1})) + x \right| \leq \delta + \frac{1}{\log n} \right) \\ \leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(|\operatorname{Re}(Z_1) + x| \leq \delta + \frac{1}{\log n} \right) \\ = \sup_{x \in \mathbb{R}} \mathbb{P} \left(|Z' + x| \leq \sqrt{2}\delta + \frac{\sqrt{2}}{\log n} \right), \end{aligned}$$

where Z' is a real standard normal random variable.

Hence, in either case, we obtain

$$\begin{aligned} \lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left\{ \left| \sum_{m=0}^N z^m (\xi_{m+1}^{(n)} + \alpha_{n,m+1}) \right| \leq \delta \right\} \cap E_n \right) \\ \leq \lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} \mathbb{P} \left(|Z' + x| \leq \sqrt{2}\delta + \frac{\sqrt{2}}{\log n} \right) \\ \leq \lim_{\delta \searrow 0} \sup_{x \in \mathbb{R}} \mathbb{P} (|Z' + x| \leq \delta) = 0 \end{aligned}$$

by a simple calculation involving the density of the standard normal distribution. This verifies (4), and the proof of Theorem 9 is complete.

3.2. Proof of Corollary 13. Let $\theta_1, \theta_2, \dots$ be a sequence of iid random variables distributed uniformly on $[0, 2\pi)$. For each $n \geq 1$, set

$$p_{2n}(z) := \prod_{j=1}^n (z - e^{i\theta_j})(z - e^{-i\theta_j})$$

and

$$p_{2n-1}(z) := (z - e^{i\theta_n}) \prod_{j=1}^{n-1} (z - e^{i\theta_j})(z - e^{-i\theta_j}).$$

We will apply Theorem 12 to show that μ'_{p_n} converges in probability to the uniform probability distribution on the unit circle centered at the origin as $n \rightarrow \infty$. From this, the conclusion of Corollary 13 follows immediately.

Define the triangular array $(\theta_j^{(n)})_{j \leq n}$ of random variables on $[0, 2\pi)$ by

$$\theta_j^{(n)} := \begin{cases} 2\pi - \theta_{j/2}, & \text{if } j \text{ even,} \\ \theta_{(j+1)/2}, & \text{if } j \text{ odd.} \end{cases}$$

It follows that, for all $n \geq 1$,

$$p_n(z) = \prod_{j=1}^n (z - e^{i\theta_j^{(n)}}).$$

Thus, it remains to show that the triangular array $(\theta_j^{(n)})_{j \leq n}$ satisfies the three assumptions of Theorem 12.

We begin by verifying condition (iii) of Theorem 12. We observe that, for any integer $m \geq 1$,

$$\frac{1}{n} \sum_{j=1}^n e^{i\theta_j^{(n)}m} = \frac{1}{n} \sum_{\substack{1 \leq j \leq n \\ j \text{ even}}} e^{i\theta_j^{(n)}m} + \frac{1}{n} \sum_{\substack{1 \leq j \leq n \\ j \text{ odd}}} e^{i\theta_j^{(n)}m}.$$

Since both sums on the right-hand side are sums of iid random variables, we apply the law of large numbers twice to obtain

$$\frac{1}{n} \sum_{j=1}^n e^{i\theta_j^{(n)}m} \longrightarrow 0$$

almost surely as $n \rightarrow \infty$.

Conditions (i) and (ii) of Theorem 12 will follow from Lemma 18 below.

Lemma 18. *Let $f : [0, 2\pi) \rightarrow \mathbb{R}$ be a function such that $f(\theta_1) + f(2\pi - \theta_1)$ is non-degenerate. Then, for any $\delta > 0$,*

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n f(\theta_j^{(n)}) \right| \leq \delta \right) = 0.$$

Proof. We consider two cases. First, if n is even, by [31, Theorem 2.22], we have

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=1}^n f(\theta_j^{(n)}) \right| \leq \delta \right) &\leq \mathbb{P} \left(\left| \sum_{j=1}^{n/2} (f(\theta_j) + f(2\pi - \theta_j)) \right| \leq \delta \right) \\ &\leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(x \leq \sum_{j=1}^{n/2} (f(\theta_j) + f(2\pi - \theta_j)) \leq x + 2\delta \right) \\ &\leq C_f \frac{1 + 2\delta}{n^{1/2}}, \end{aligned}$$

where $C_f > 0$ is a constant that only depends on f . In the case that $n > 1$ is odd, by [31, Lemma 1.11] and [31, Theorem 2.22], we obtain

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=1}^n f(\theta_j^{(n)}) \right| \leq \delta \right) &\leq \mathbb{P} \left(\left| \sum_{j=1}^{(n-1)/2} (f(\theta_j) + f(2\pi - \theta_j)) + f(\theta_n^{(n)}) \right| \leq \delta \right) \\ &\leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(x \leq \sum_{j=1}^{(n-1)/2} (f(\theta_j) + f(2\pi - \theta_j)) + f(\theta_n^{(n)}) \leq x + 2\delta \right) \\ &\leq \sup_{x \in \mathbb{R}} \mathbb{P} \left(x \leq \sum_{j=1}^{(n-1)/2} (f(\theta_j) + f(2\pi - \theta_j)) \leq x + 2\delta \right) \\ &\leq C_f \frac{1 + 2\delta}{(n-1)^{1/2}}, \end{aligned}$$

and the proof is complete. \square

To verify condition (i) of Theorem 12, we note that, for any $\delta > 0$,

$$\mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) \leq \mathbb{P} \left(\left| \sum_{j=1}^n \cos(\theta_j^{(n)}) \right| \leq \delta \right).$$

Thus, by Lemma 18, we conclude that, for any $\delta > 0$,

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) = 0.$$

It remains to verify condition (ii) of Theorem 12. To this end, fix $z \in \mathbb{D}$, and let $\eta > 0$. As

$$\left| \sum_{j=1}^n \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| \leq \sum_{j=1}^n \left| \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| \leq \frac{n}{1 - |z|},$$

it follows that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{n} \log \left| \sum_{j=1}^n \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| > \eta \right) = 0.$$

On the other hand,

$$\begin{aligned} \mathbb{P} \left(\frac{1}{n} \log \left| \sum_{j=1}^n \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| < -\eta \right) &\leq \mathbb{P} \left(\left| \sum_{j=1}^n \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| < e^{-n\eta} \right) \\ &\leq \mathbb{P} \left(\left| \sum_{j=1}^n \frac{\operatorname{Re}(z) - \cos(\theta_j^{(n)})}{|z - e^{-i\theta_j^{(n)}}|^2} \right| < e^{-n\eta} \right), \end{aligned}$$

and hence

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{n} \log \left| \sum_{j=1}^n \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| < -\eta \right) = 0$$

by Lemma 18. Therefore, we conclude that

$$\frac{1}{n} \log \left| \sum_{j=1}^n \frac{1}{z - e^{-i\theta_j^{(n)}}} \right| \rightarrow 0$$

in probability as $n \rightarrow \infty$, and the proof of Corollary 13 is complete.

4. PROOF OF THEOREMS 11 AND 12

This section is devoted to the proof of Theorems 11 and 12.

4.1. Convergence of radial components implies convergence of the empirical measures. Both Theorems 11 and 12 will follow from Lemma 19 below.

Lemma 19 (Convergence of radial components implies convergence of measures). *For each $n \geq 1$, let $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ be random variables on $[0, 2\pi)$, and set*

$$p_n(z) := \prod_{j=1}^n (z - e^{i\theta_j^{(n)}}).$$

Let $r_1^{(n)} e^{i\phi_1^{(n)}}, \dots, r_{n-1}^{(n)} e^{i\phi_{n-1}^{(n)}}$ be the zeros of p'_n in polar form. Assume

(i) *for all integers $m \geq 1$,*

$$\frac{1}{n} \sum_{j=1}^n e^{im\theta_j^{(n)}} \rightarrow 0 \tag{5}$$

in probability as $n \rightarrow \infty$,

(ii) *we have*

$$\frac{1}{n-1} \sum_{j=1}^{n-1} (1 - r_j^{(n)}) \rightarrow 0 \tag{6}$$

in probability as $n \rightarrow \infty$.

Then μ'_{p_n} converges in probability as $n \rightarrow \infty$ to the uniform probability distribution on the unit circle centered at the origin.

Remark 20. By the Gauss–Lucas theorem (Theorem 1), it follows that

$$\sup_{n \geq 1} \max_{1 \leq j \leq n} r_j^{(n)} \leq 1.$$

Thus, condition (ii) of Lemma 19 implies that most of the roots of p'_n are close to the unit circle centered at the origin.

The remainder of this subsection will be devoted to proving Lemma 19. In particular, we will need the following result, which is adapted from [40, Proposition 3.2].

Lemma 21. *Let $n \geq 2$. Let $x_1, \dots, x_n \in \mathbb{C}$ with $|x_j| \leq \tau$ for all $1 \leq j \leq n$. Let y_1, \dots, y_{n-1} be the critical points of $p(z) := \prod_{j=1}^n (z - x_j)$. Then, for any integer $k \geq 1$, there exists a constant $C > 0$ (depending only on τ and k) such that*

$$\left| \frac{1}{n} \sum_{j=1}^n x_j^k - \frac{1}{n-1} \sum_{j=1}^{n-1} y_j^k \right| \leq \frac{C}{n-1}.$$

In order to prove Lemma 21, we will need the following result from [4].

Lemma 22. *Let $n \geq 2$. If $x_1, \dots, x_n \in \mathbb{C}$ are the roots of $p(z) := \prod_{j=1}^n (z - x_j)$, and p has critical points y_1, \dots, y_{n-1} , then the matrix*

$$D \left(I_{n-1} - \frac{1}{n} J \right) + \frac{x_n}{n} J$$

has y_1, \dots, y_{n-1} as its eigenvalues, where $D = \text{diag}(x_1, \dots, x_{n-1})$, I_{n-1} is the identity matrix of order $n-1$, and J is the $(n-1) \times (n-1)$ matrix of all entries 1.

We now prove Lemma 21.

Proof of Lemma 21. The proof presented here is adapted from the proof given in [40]. We observe that it suffices to show

$$\left| \sum_{j=1}^{n-1} x_j^k - \sum_{j=1}^{n-1} y_j^k \right| \leq C,$$

where $C > 0$ depends only on τ and k .

Let $D = \text{diag}(x_1, \dots, x_{n-1})$. Then, by Lemma 22, it follows that

$$\sum_{j=1}^{n-1} y_j^k = \text{tr} \left(D - \frac{1}{n} D J - \frac{x_n}{n} J \right)^k,$$

where J is the $(n-1) \times (n-1)$ matrix of all entries 1. Thus, it suffices to show

$$\left| \text{tr} \left(D - \frac{1}{n} D J - \frac{x_n}{n} J \right)^k - \text{tr} D^k \right| \leq C. \quad (7)$$

We note that $\left(D - \frac{1}{n} D J - \frac{x_n}{n} J \right)^k$ can be written as the sum over all terms of the form

$$D^{l_1} \left(-\frac{1}{n} D J \right)^{l_2} \left(\frac{x_n}{n} J \right)^{l_3} \dots D^{l_{3k-2}} \left(-\frac{1}{n} D J \right)^{l_{3k-1}} \left(\frac{x_n}{n} J \right)^{l_{3k}}, \quad (8)$$

where l_1, \dots, l_{3k} are non-negative integers such that $l_{3j-2} + l_{3j-1} + l_{3j} = 1$ for each $1 \leq j \leq k$. The total number of such terms is 3^k . One of the terms is D^k . We will show that each of the remaining $3^k - 1$ terms can be uniformly bounded by a constant which only depends on τ and k .

Fix l_1, \dots, l_{3k} such that the term given in (8) is not D^k . In order to simplify the expression in (8), we observe that

$$J^m = (n-1)^{m-1} J$$

for all $m \geq 1$. We also have

$$(D^p J)(D^q J) = \left(\sum_{j=1}^{n-1} x_j^q \right) (D^p J),$$

for any $p, q \geq 0$.

Thus, the term in (8) can be written as

$$(-1)^p x_n^q \left(\frac{n-1}{n} \right)^{s_0} \left(\frac{\sum_{j=1}^{n-1} x_j}{n} \right)^{s_1} \dots \left(\frac{\sum_{j=1}^{n-1} x_j^{k-1}}{n} \right)^{s_{k-1}} M, \quad (9)$$

where $p, q, s_0, \dots, s_{k-1}$ are non-negative integers no larger than k , and M is a $(n-1) \times (n-1)$ matrix. In particular, M is of the form $\frac{1}{n} D^m J$ or $\frac{1}{n} D^{m_1} J D^{m_2}$ for some non-negative integers m, m_1, m_2 which are no larger than k .

The scalar term in (9) can be uniformly bounded by a constant depending only on τ and k since $\max_{1 \leq j \leq n} |x_j| \leq \tau$. If $M = \frac{1}{n} D^m J$, then

$$|\text{tr}(M)| = \frac{1}{n} \left| \sum_{j=1}^{n-1} x_j^m \right| \leq \frac{n-1}{n} \tau^m \leq \tau^k$$

since $m \leq k$. Similarly, if $M = \frac{1}{n} D^{m_1} J D^{m_2}$, then

$$|\text{tr}(M)| = \frac{1}{n} |\text{tr}(D^{m_1+m_2} J)| = \frac{1}{n} \left| \sum_{j=1}^{n-1} x_j^{m_1+m_2} \right| \leq \tau^{2k}$$

because $m_1, m_2 \leq k$.

Combining the bounds above yields (7), and the proof is complete. \square

With Lemma 21 in hand, we can now prove Lemma 19.

Proof of Lemma 19. By (5) and Lemma 21, it follows that, for each $m \geq 1$,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} \left(r_j^{(n)} \right)^m e^{i\phi_j^{(n)} m} \rightarrow 0$$

in probability as $n \rightarrow \infty$. By the Gauss–Lucas theorem (Theorem 1),

$$\sup_{n \geq 1} \max_{1 \leq j \leq n} r_j^{(n)} \leq 1.$$

Thus,

$$\begin{aligned} \frac{1}{n-1} \left| \sum_{j=1}^{n-1} \left(r_j^{(n)} \right)^m e^{i\phi_j^{(n)} m} - \sum_{j=1}^{n-1} e^{i\phi_j^{(n)} m} \right| &\leq \frac{1}{n-1} \sum_{j=1}^{n-1} \left| 1 - \left(r_j^{(n)} \right)^m \right| \\ &\leq \frac{C_m}{n-1} \sum_{j=1}^{n-1} \left(1 - r_j^{(n)} \right), \end{aligned}$$

where $C_m > 0$ depends only on m . Hence, by (6), we conclude that, for any $m \geq 1$,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} e^{im\phi_j^{(n)}} \rightarrow 0$$

in probability as $n \rightarrow \infty$. This also implies that, for any $m \geq 1$,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} e^{-im\phi_j^{(n)}} \rightarrow 0$$

in probability as $n \rightarrow \infty$. In other words, for any trigonometric polynomial q ,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} q(\phi_j^{(n)}) \rightarrow \mathbb{E}[q(\xi)] \quad (10)$$

in probability as $n \rightarrow \infty$, where ξ is a random variable uniformly distributed on $[0, 2\pi)$.

Let $f : \mathbb{C} \rightarrow \mathbb{R}$ be a bounded Lipschitz continuous function. By the Portemanteau theorem (see, for example, [21, Theorem 13.16]), it suffices to show that

$$\frac{1}{n-1} \sum_{j=1}^{n-1} f\left(r_j^{(n)} e^{i\phi_j^{(n)}}\right) \rightarrow \mathbb{E}[f(e^{i\xi})]$$

in probability as $n \rightarrow \infty$.

Let $\varepsilon > 0$. By [33, Theorem 4.25], there exists a trigonometric polynomial q such that

$$\sup_{t \in [0, 2\pi]} |f(e^{it}) - q(t)| \leq \varepsilon. \quad (11)$$

Then, by the triangle inequality, we have

$$\begin{aligned} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} f\left(r_j^{(n)} e^{i\phi_j^{(n)}}\right) - \mathbb{E}[f(e^{i\xi})] \right| &\leq \frac{1}{n-1} \left| \sum_{j=1}^{n-1} f\left(r_j^{(n)} e^{i\phi_j^{(n)}}\right) - \sum_{j=1}^{n-1} f\left(e^{i\phi_j^{(n)}}\right) \right| \\ &\quad + \frac{1}{n-1} \left| \sum_{j=1}^{n-1} f\left(e^{i\phi_j^{(n)}}\right) - \sum_{j=1}^{n-1} q\left(\phi_j^{(n)}\right) \right| \\ &\quad + \left| \frac{1}{n-1} \sum_{j=1}^{n-1} q\left(\phi_j^{(n)}\right) - \mathbb{E}[q(\xi)] \right| \\ &\quad + \left| \mathbb{E}[q(\xi)] - \mathbb{E}[f(e^{i\xi})] \right|. \end{aligned}$$

Since f is Lipschitz continuous, we obtain

$$\frac{1}{n-1} \left| \sum_{j=1}^{n-1} f\left(r_j^{(n)} e^{i\phi_j^{(n)}}\right) - \sum_{j=1}^{n-1} f\left(e^{i\phi_j^{(n)}}\right) \right| \leq \frac{C_f}{n-1} \sum_{j=1}^{n-1} (1 - r_j^{(n)}),$$

where C_f is the Lipschitz constant of f . By (11), we have

$$\frac{1}{n-1} \left| \sum_{j=1}^{n-1} f\left(e^{i\phi_j^{(n)}}\right) - \sum_{j=1}^{n-1} q\left(\phi_j^{(n)}\right) \right| \leq \varepsilon$$

and

$$|\mathbb{E}[q(\xi)] - \mathbb{E}[f(e^{i\xi})]| \leq \varepsilon.$$

Thus, we conclude that

$$\begin{aligned} & \left| \frac{1}{n-1} \sum_{j=1}^{n-1} f\left(r_j^{(n)} e^{i\phi_j^{(n)}}\right) - \mathbb{E}[f(e^{i\xi})] \right| \\ & \leq \frac{C_f}{n-1} \sum_{j=1}^{n-1} (1 - r_j^{(n)}) + \left| \frac{1}{n-1} \sum_{j=1}^{n-1} q\left(\phi_j^{(n)}\right) - \mathbb{E}[q(\xi)] \right| + 2\varepsilon. \end{aligned}$$

The claim now follows from (6) and (10). \square

4.2. Convergence of the radial components. In order to apply Lemma 19, we must verify the convergence in (6). We do so in the following lemmata.

Lemma 23. *For each $n \geq 1$, let $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ be random variables on $[0, 2\pi)$, and set*

$$p_n(z) := \prod_{j=1}^n (z - e^{i\theta_j^{(n)}}).$$

Let $\zeta_1^{(n)}, \dots, \zeta_{n-1}^{(n)}$ be the zeros of p'_n . Assume

(i) we have

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) = 0,$$

(ii) for almost every $z \in \mathbb{D}$,

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{m=0}^{\lfloor \log^2 n \rfloor} z^m \sum_{j=1}^n e^{-i\theta_j^{(n)}(m+1)} \right| \leq \delta \right) = 0.$$

Then, for any $0 < \varepsilon < 1$ and for every infinitely differentiable function $\varphi : \mathbb{C} \rightarrow \mathbb{R}$ supported on $\mathbb{D}_{1-\varepsilon}$,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} \varphi(\zeta_j^{(n)}) \rightarrow 0$$

in probability as $n \rightarrow \infty$.

We also have the following alternative formulation of Lemma 23.

Lemma 24 (Alternative formulation). *For each $n \geq 1$, let $\theta_1^{(n)}, \dots, \theta_n^{(n)}$ be random variables on $[0, 2\pi)$, and set*

$$p_n(z) := \prod_{j=1}^n (z - e^{i\theta_j^{(n)}}).$$

Let $\zeta_1^{(n)}, \dots, \zeta_{n-1}^{(n)}$ be the zeros of p'_n . Assume

(i) we have

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) = 0,$$

(ii) for almost every $z \in \mathbb{D}$,

$$\frac{1}{n} \log \left| \sum_{j=1}^n \frac{1}{z - e^{i\theta_j^{(n)}}} \right| \rightarrow 0$$

in probability as $n \rightarrow \infty$.

Then, for any $0 < \varepsilon < 1$ and for every infinitely differentiable function $\varphi : \mathbb{C} \rightarrow \mathbb{R}$ supported on $\mathbb{D}_{1-\varepsilon}$,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} \varphi(\zeta_j^{(n)}) \rightarrow 0$$

in probability as $n \rightarrow \infty$.

We will prove Lemmas 23 and 24 in Section 5. We now complete the proof of Theorems 11 and 12 assuming Lemmas 23 and 24. We prove both theorems simultaneously.

Proof of Theorems 11 and 12. Let $\zeta_1^{(n)}, \dots, \zeta_{n-1}^{(n)}$ be the zeros of p'_n . In view of Lemma 19, it suffices to show that

$$\frac{1}{n-1} \sum_{j=1}^{n-1} (1 - |\zeta_j^{(n)}|) \rightarrow 0$$

in probability as $n \rightarrow \infty$.

Let $0 < \varepsilon < 1/2$. Let $\varphi : \mathbb{C} \rightarrow [0, 1]$ be an infinitely differentiable function such that φ takes the value 1 on $\mathbb{D}_{1-2\varepsilon}$ and takes the value zero on $\mathbb{C} \setminus \mathbb{D}_{1-\varepsilon}$. By Lemma 23 (alternatively, Lemma 24), we have

$$\frac{1}{n-1} \sum_{j=1}^{n-1} \varphi(\zeta_j^{(n)}) \rightarrow 0 \tag{12}$$

in probability as $n \rightarrow \infty$.

On the other hand, by the Gauss–Lucas theorem (Theorem 1), it follows that

$$\sup_{n \geq 1} \max_{1 \leq j \leq n} |\zeta_j^{(n)}| \leq 1.$$

Thus, we have

$$\begin{aligned} \frac{1}{n-1} \sum_{j=1}^{n-1} (1 - |\zeta_j^{(n)}|) &= \frac{1}{n-1} \sum_{j=1}^{n-1} (1 - |\zeta_j^{(n)}|) \mathbf{1}_{\{\zeta_j^{(n)} \in \mathbb{D}_{1-2\varepsilon}\}} \\ &\quad + \frac{1}{n-1} \sum_{j=1}^{n-1} (1 - |\zeta_j^{(n)}|) \mathbf{1}_{\{\zeta_j^{(n)} \notin \mathbb{D}_{1-2\varepsilon}\}} \\ &\leq \frac{1}{n-1} \sum_{j=1}^{n-1} \mathbf{1}_{\{\zeta_j^{(n)} \in \mathbb{D}_{1-2\varepsilon}\}} + 2\varepsilon \\ &\leq \frac{1}{n-1} \sum_{j=1}^{n-1} \varphi(\zeta_j^{(n)}) + 2\varepsilon. \end{aligned}$$

Since ε was arbitrary, the claim now follows from (12). \square

5. PROOF OF LEMMAS 23 AND 24

It remains to verify Lemmas 23 and 24. The proof is based on a connection with logarithmic potential theory. In particular, we will exploit the following formula from [15, Section 2.4.1]: for every analytic function f which does not vanish identically,

$$\frac{1}{2\pi} \Delta \log |f| = \sum_{z \in \mathbb{C}: f(z)=0} \mathcal{N}_f(z) \delta_z, \quad (13)$$

where $\mathcal{N}_f(z)$ is the multiplicity of the zero at z and δ_z is the unit point mass at z . Here Δ is the Laplace operator, which should be interpreted in the distributional sense. Similar methods also appeared in [19, 20, 41]. In fact, our overall strategy is based on the arguments presented in [19].

Let $\theta_1^{(n)}, \dots, \theta_n^{(n)}, \zeta_1^{(n)}, \dots, \zeta_{n-1}^{(n)}$, and p_n be as in Lemma 23 (alternatively, Lemma 24). Consider the logarithmic derivative of p_n :

$$L_n(z) := \frac{p'_n(z)}{p_n(z)} = \sum_{j=1}^n \frac{1}{z - e^{i\theta_j^{(n)}}}. \quad (14)$$

Let $0 < \varepsilon < 1$, and let $\varphi : \mathbb{C} \rightarrow \mathbb{R}$ be an infinitely differentiable function supported on $\mathbb{D}_{1-\varepsilon}$. In view of (13), we have

$$\frac{1}{2\pi n} \int_{\mathbb{C}} (\log |L_n(z)|) \Delta \varphi(z) d\lambda(z) = \frac{1}{n} \sum_{j=1}^{n-1} \varphi(\zeta_j^{(n)}) - \frac{1}{n} \sum_{j=1}^n \varphi(e^{i\theta_j^{(n)}}),$$

where λ denotes Lebesgue measure on \mathbb{C} . Since φ is supported on $\mathbb{D}_{1-\varepsilon}$, the above equality becomes

$$\frac{1}{2\pi n} \int_{\mathbb{C}} (\log |L_n(z)|) \Delta \varphi(z) d\lambda(z) = \frac{1}{n} \sum_{j=1}^{n-1} \varphi(\zeta_j^{(n)}).$$

Therefore, the proof of Lemma 23 (alternatively, Lemma 24) reduces to showing that

$$\frac{1}{n} \int_{\mathbb{C}} (\log |L_n(z)|) \Delta \varphi(z) d\lambda(z) \longrightarrow 0 \quad (15)$$

in probability as $n \rightarrow \infty$.

In order to verify (15), we will need the following result from [41].

Lemma 25 (Lemma 3.1 from [41]). *Let (X, \mathcal{A}, ν) be a finite measure space. Let $f_1, f_2, \dots : X \rightarrow \mathbb{R}$ be random functions which are defined over a probability space $(\Omega, \mathcal{B}, \mathbb{P})$ and are jointly measurable with respect to $\mathcal{A} \otimes \mathcal{B}$. Assume that:*

- (i) *for ν -almost every $x \in X$, $f_n(x)$ converges in probability to zero as $n \rightarrow \infty$,*
- (ii) *for some $\delta > 0$, the sequence $\int_X |f_n(x)|^{1+\delta} d\nu(x)$ is tight.*

Then $\int_X f_n(x) d\nu(x)$ converges in probability to zero as $n \rightarrow \infty$.

In order to apply Lemma 25, we will show that $\frac{1}{n} \log |L_n(z)|$ converges in probability to zero for a.e. $z \in \mathbb{D}_{1-\varepsilon}$ and that the sequence $\frac{1}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |L_n(z)| d\lambda(z)$ is tight. To this end, we define

$$\log_- x := \begin{cases} |\log x|, & 0 \leq x \leq 1 \\ 0, & x > 1, \end{cases} \quad \log_+ x := \begin{cases} 0 & 0 \leq x \leq 1 \\ \log x, & x > 1. \end{cases}$$

From (14) it follows that $L_n(z)$ is finite for all $z \in \mathbb{D}$. Moreover, $L_n(z) = 0$ only when $p'_n(z) = 0$; in this case, $\log_- |L_n(z)| = \infty$.

5.1. Pointwise convergence of $L_n(z)$. This subsection is devoted to the following lemma.

Lemma 26. *If, for almost every $z \in \mathbb{D}$,*

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{m=0}^{\lfloor \log^2 n \rfloor} z^m \sum_{j=1}^n e^{-i\theta_j^{(n)}(m+1)} \right| \leq \delta \right) = 0, \quad (16)$$

then, for almost every $z \in \mathbb{D}$,

$$\frac{1}{n} \log |L_n(z)| \longrightarrow 0$$

in probability as $n \rightarrow \infty$.

Proof. For $|z| < 1$, we have, by Fubini's theorem,

$$\begin{aligned} L_n(z) &= - \sum_{j=1}^n \frac{1}{e^{i\theta_j^{(n)}}} \frac{1}{1 - \frac{z}{e^{i\theta_j^{(n)}}}} = - \sum_{j=1}^n \frac{1}{e^{i\theta_j^{(n)}}} \sum_{m=0}^{\infty} \frac{z^m}{e^{i\theta_j^{(n)}m}} = - \sum_{m=0}^{\infty} z^m \sum_{j=1}^n e^{-i\theta_j^{(n)}(m+1)} \\ &= - \sum_{m=0}^{\infty} z^m T_m^{(n)}(z), \end{aligned}$$

where

$$T_m^{(n)}(z) := \sum_{j=1}^n e^{-i\theta_j^{(n)}(m+1)}.$$

Here the use of Fubini's theorem is justified since

$$\sum_{m=0}^{\infty} \sum_{j=1}^n |z|^m \left| e^{-i\theta_j^{(n)}(m+1)} \right| \leq n \sum_{m=0}^{\infty} |z|^m < \infty$$

for all $|z| < 1$ and every $n \geq 1$.

Let $0 < \varepsilon < 1$, and fix $z \in \mathbb{D}$ with $|z| \leq 1 - \varepsilon$ such that (16) holds. Set $N := \lfloor \log^2 n \rfloor$. We can then write

$$|L_n(z)| = \left| \sum_{m=0}^N z^m T_m^{(n)}(z) + \sum_{m=N+1}^{\infty} z^m T_m^{(n)}(z) \right|. \quad (17)$$

Since $|T_m^{(n)}(z)| \leq n$ for all integers $m \geq 0$, it follows that

$$\left| \sum_{m=N+1}^{\infty} z^m T_m^{(n)}(z) \right| \leq \frac{n|z|^{N+1}}{1 - |z|} \leq \frac{n(1 - \varepsilon)^{N+1}}{\varepsilon}. \quad (18)$$

In addition, we observe that

$$|L_n(z)| \leq \sum_{j=1}^n \left| \frac{1}{z - e^{i\theta_j^{(n)}}} \right| \leq \frac{n}{\varepsilon}. \quad (19)$$

Since ε is arbitrary, it suffices to show that $\frac{1}{n} \log |L_n(z)|$ converges to zero in probability. Since

$$\log |L_n(z)| = \log_+ |L_n(z)| - \log_- |L_n(z)|,$$

it suffices to show that both $\frac{1}{n} \log_+ |L_n(z)|$ and $\frac{1}{n} \log_- |L_n(z)|$ converge to zero in probability.

For the first term, we have

$$0 \leq \frac{1}{n} \log_+ |L_n(z)| \leq \frac{1}{n} \log \left(\frac{n}{\varepsilon} \right)$$

by (19). Thus, $\frac{1}{n} \log_+ |L_n(z)|$ converges to zero a.s.

It now suffices to show that, for any $\eta > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{n} \log_- |L_n(z)| > \eta \right) = 0.$$

However, from (17) and (18), we observe that

$$\begin{aligned} \log_- |L_n(z)| &\leq \frac{1}{|L_n(z)|} \\ &= \frac{1}{\left| \sum_{m=0}^N z^m T_m^{(n)}(z) + \sum_{m=N+1}^{\infty} z^m T_m^{(n)}(z) \right|} \\ &\leq \frac{1}{\left| \sum_{m=0}^N z^m T_m^{(n)}(z) \right| - \frac{n(1-\varepsilon)^{N+1}}{\varepsilon}} \end{aligned}$$

provided

$$\left| \sum_{m=0}^N z^m T_m^{(n)}(z) \right| > \frac{n(1-\varepsilon)^{N+1}}{\varepsilon}.$$

Thus, we obtain

$$\begin{aligned} &\mathbb{P} \left(\frac{1}{n} \log_- |L_n(z)| > \eta \right) \\ &\leq \mathbb{P} \left(\frac{1}{n} \log_- |L_n(z)| > \eta \text{ and } \left| \sum_{m=0}^N z^m T_m^{(n)}(z) \right| > \frac{n(1-\varepsilon)^{N+1}}{\varepsilon} \right) \\ &\quad + \mathbb{P} \left(\left| \sum_{m=0}^N z^m T_m^{(n)}(z) \right| \leq \frac{n(1-\varepsilon)^{N+1}}{\varepsilon} \right) \\ &\leq 2\mathbb{P} \left(\left| \sum_{m=0}^N z^m T_m^{(n)}(z) \right| \leq \frac{1}{n\eta} + \frac{n(1-\varepsilon)^{N+1}}{\varepsilon} \right). \end{aligned}$$

Since

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n\eta} + \frac{n(1-\varepsilon)^{N+1}}{\varepsilon} \right) = 0,$$

the claim now follows from (16). \square

5.2. Tightness. This subsection is devoted to proving the following lemma.

Lemma 27. *If*

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq \delta \right) = 0, \quad (20)$$

then, for any $0 < \varepsilon < 1$, the sequence $\frac{1}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |L_n(z)| d\lambda(z)$ is tight.

The proof of Lemma 27 is based on the arguments presented in [19].

Proof of Lemma 27. Let $0 < \varepsilon < 1$. Define $R := 1 - \varepsilon/2$. Note that $L_n(z)$ has no poles in the closed disk $\overline{\mathbb{D}}_R$. Let $\zeta_1^{(n)}, \dots, \zeta_{k_n}^{(n)}$ denote the zeros of L_n in \mathbb{D}_R , where $k_n \leq n$. So by the Poisson–Jensen formula (see, for instance, [26, Chapter II.8]), for $z := re^{i\theta} \in \mathbb{D}_{1-\varepsilon}$ other than a zero, we have

$$\log |L_n(z)| = I_n(z) + \sum_{l=1}^{k_n} \log \left| \frac{R(z - \zeta_l^{(n)})}{R^2 - \bar{\zeta}_l^{(n)} z} \right|, \quad (21)$$

where

$$I_n(z) := \frac{1}{2\pi} \int_0^{2\pi} \log |L_n(Re^{i\phi})| P(z, Re^{i\phi}) d\phi \quad (22)$$

and

$$P(z, Re^{i\phi}) := \frac{R^2 - r^2}{R^2 + r^2 - 2Rr \cos(\theta - \phi)}, \quad r < R. \quad (23)$$

Thus, by the Cauchy–Schwarz inequality, we have

$$\begin{aligned} (\log |L_n(z)|)^2 &\leq 2|I_n(z)|^2 + 2 \left(\sum_{l=1}^{k_n} \log \left| \frac{R(z - \zeta_l^{(n)})}{R^2 - \bar{\zeta}_l^{(n)} z} \right| \right)^2 \\ &\leq 2|I_n(z)|^2 + 2k_n \sum_{l=1}^{k_n} \log^2 \left| \frac{R(z - \zeta_l^{(n)})}{R^2 - \bar{\zeta}_l^{(n)} z} \right|. \end{aligned}$$

Hence, we conclude that

$$\begin{aligned} \frac{1}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |L_n(z)| d\lambda(z) &\leq \frac{2}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} |I_n(z)|^2 d\lambda(z) \\ &\quad + \frac{2k_n}{n^2} \sum_{l=1}^{k_n} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 \left| \frac{R(z - \zeta_l^{(n)})}{R^2 - \bar{\zeta}_l^{(n)} z} \right| d\lambda(z). \end{aligned} \quad (24)$$

Observe that

$$\inf_{z \in \mathbb{D}_{1-\varepsilon}} \inf_{y \in \mathbb{D}_R} |R^2 - yz| \geq C_1,$$

where $C_1 > 0$ depends only on ε . Similarly,

$$\sup_{z \in \mathbb{D}_{1-\varepsilon}} \sup_{y \in \mathbb{D}_R} |R^2 - yz| \leq 2.$$

Thus, for any $y \in \mathbb{D}_R$, we obtain

$$\begin{aligned}
\int_{\mathbb{D}_{1-\varepsilon}} \log^2 \left| \frac{R(z-y)}{R^2-yz} \right| d\lambda(z) &\leq 2 \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |R(z-y)| d\lambda(z) \\
&\quad + 2 \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |R^2-yz| d\lambda(z) \\
&\leq 4\pi \log^2 |R| + 4 \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |z-y| d\lambda(z) \\
&\quad + 4 \int_{\mathbb{D}_{1-\varepsilon}} \left(\frac{1}{|R^2-yz|^2} + |R^2-yz|^2 \right) d\lambda(z) \\
&\leq 4\pi \log^2 |R| + 4 \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |z-y| d\lambda(z) \\
&\quad + 4\pi \left(\frac{1}{C_1^2} + 4 \right) \\
&\leq C_2,
\end{aligned}$$

where $C_2 > 0$ depends only on ε . Here we used that $\log |\cdot|$ is square integrable as well as the bound $\log^2 |x| \leq \frac{2}{|x|^2} + 2|x|^2$. Therefore, we conclude that

$$\sup_{y \in \mathbb{D}_R} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 \left| \frac{R(z-y)}{R^2-yz} \right| d\lambda(z) \leq C_2,$$

and hence (since $k_n \leq n$)

$$\frac{k_n}{n^2} \sum_{l=1}^{k_n} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 \left| \frac{R(z-\zeta_l^{(n)})}{R^2-\bar{\zeta}_l^{(n)}z} \right| d\lambda(z) \leq C_2.$$

Thus, in view of (24), it suffices to show that

$$\frac{1}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} |I_n(z)|^2 d\lambda(z)$$

is tight.

We recall that, for $z = re^{i\theta}$,

$$I_n(z) = \frac{1}{2\pi} \int_0^{2\pi} \log |L_n(Re^{i\phi})| \frac{R^2 - r^2}{R^2 + r^2 - 2Rr \cos(\theta - \phi)} d\phi.$$

We now observe that, for any $\theta, \phi \in [0, 2\pi)$ and every $0 \leq r \leq 1 - \varepsilon$, we have

$$C'_3 \leq R^2 - r^2 \leq C_3$$

and

$$C'_4 \geq R^2 + r^2 - 2Rr \cos(\theta - \phi) \geq (R - r)^2 \geq C_4,$$

where $C_3, C'_3, C_4, C'_4 > 0$ depend only on ε .

We write

$$\begin{aligned}
I_n(z) &= \frac{1}{2\pi} \int_0^{2\pi} \log_+ |L_n(Re^{i\phi})| P(z, Re^{i\phi}) d\phi \\
&\quad - \frac{1}{2\pi} \int_0^{2\pi} \log_- |L_n(Re^{i\phi})| P(z, Re^{i\phi}) d\phi.
\end{aligned}$$

In particular,

$$\begin{aligned} I_n(z) &\leq \frac{1}{2\pi} \int_0^{2\pi} \log_+ |L_n(Re^{i\phi})| P(z, Re^{i\phi}) d\phi \\ &\leq \frac{C_3}{C_4} \frac{1}{2\pi} \int_0^{2\pi} \log_+ |L_n(Re^{i\phi})| d\phi. \end{aligned}$$

On the other hand,

$$\begin{aligned} I_n(z) &\geq \frac{C'_3}{C'_4} \frac{1}{2\pi} \int_0^{2\pi} \log_+ |L_n(Re^{i\phi})| d\phi - \frac{C_3}{C_4} \frac{1}{2\pi} \int_0^{2\pi} \log_- |L_n(Re^{i\phi})| d\phi \\ &= \frac{C_3}{C_4} \frac{1}{2\pi} \int_0^{2\pi} \log |L_n(Re^{i\phi})| d\phi + \left(\frac{C'_3}{C'_4} - \frac{C_3}{C_4} \right) \frac{1}{2\pi} \int_0^{2\pi} \log_+ |L_n(Re^{i\phi})| d\phi \\ &\geq \frac{C_3}{C_4} I_n(0) - \left| \frac{C'_3}{C'_4} - \frac{C_3}{C_4} \right| \frac{1}{2\pi} \int_0^{2\pi} \log_+ |L_n(Re^{i\phi})| d\phi \end{aligned}$$

by definition of $I_n(0)$ (see (22) and (23)). Since

$$\log_+ |L_n(Re^{i\phi})| \leq \log_+ \left(\sum_{j=1}^n \left| \frac{1}{e^{i\theta_j^{(n)}} - Re^{i\phi}} \right| \right) \leq \log \left(\frac{2n}{\varepsilon} \right)$$

uniformly in ϕ , we conclude that

$$I_n(z) \leq \frac{C_3}{C_4} \log \left(\frac{2n}{\varepsilon} \right)$$

and

$$I_n(z) \geq \frac{C_3}{C_4} I_n(0) - \left| \frac{C'_3}{C'_4} - \frac{C_3}{C_4} \right| \log \left(\frac{2n}{\varepsilon} \right)$$

for all $z \in \mathbb{D}_{1-\varepsilon}$.

As

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{2n}{\varepsilon} \right) = 0,$$

it suffices to show that $\frac{1}{n} I_n(0)$ is bounded below in probability.

From (21), we observe that

$$I_n(0) \geq \log |L_n(0)|$$

since

$$\sum_{l=1}^{k_n} \log \left| \frac{\zeta_l^{(n)}}{R} \right| \leq 0.$$

Thus, for any $\eta > 0$, we have

$$\begin{aligned} \mathbb{P} \left(\frac{1}{n} I_n(0) \leq -\eta \right) &\leq \mathbb{P} \left(\frac{1}{n} \log |L_n(0)| \leq -\eta \right) \\ &\leq \mathbb{P} (|L_n(0)| \leq e^{-\eta n}) \\ &= \mathbb{P} \left(\left| \sum_{j=1}^n e^{-i\theta_j^{(n)}} \right| \leq e^{-\eta n} \right). \end{aligned}$$

From (20), we obtain that, for any $\eta > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{n} I_n(0) \leq -\eta \right) = 0.$$

Combining the bounds above, we conclude that the sequence

$$\frac{1}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} |I_n(z)|^2 d\lambda(z)$$

is tight, and the proof is complete. \square

5.3. Completing the proof of Lemmas 23 and 24. We now complete the proof of Lemmas 23 and 24. Indeed, in view of Lemma 25, the proof reduces to showing that

- (i) for a.e. $z \in \mathbb{D}$, $\frac{1}{n} \log |L_n(z)|$ converges in probability to zero as $n \rightarrow \infty$,
- (ii) for any $0 < \varepsilon < 1$, the sequence $\frac{1}{n^2} \int_{\mathbb{D}_{1-\varepsilon}} \log^2 |L_n(z)| d\lambda(z)$ is tight.

Thus, Lemma 23 follows from Lemmas 26 and 27. Lemma 24 follows from Lemma 27 as the convergence of $\frac{1}{n} \log |L_n(z)|$ to zero is assumed in the statement of the lemma.

REFERENCES

- [1] A. Aziz, *On the zeros of a polynomial and its derivative*, Bull. Austral. Math. Soc., 31(2):245–255, 1985.
- [2] Z. D. Bai, J. Silverstein, *Spectral analysis of large dimensional random matrices*, Mathematics Monograph Series **2**, Science Press, Beijing 2006.
- [3] H. E. Bray, *On the Zeros of a Polynomial and of Its Derivative*, Amer. J. Math., 53(4):864–872, 1931.
- [4] W. S. Cheung, T. W. Ng, *A companion matrix approach to the study of zeros and critical points of a polynomial*, J. Math. Anal. Appl. 319 (2006), no. 2, 690–707.
- [5] B. Ćurgus, V. Mascioni, *A contraction of the Lucas polygon*, Proc. Amer. Math. Soc., 132(10):2973–2981 (electronic), 2004.
- [6] P. Diaconis, S. N. Evans, *Linear functionals of eigenvalues of random matrices*, Trans. Amer. Math. Soc. vol. 353, no. 7 (2001), 2615–2633.
- [7] P. Diaconis, M. Shahshahani, *On the eigenvalues of random matrices. Studies in applied probability*, J. Appl. Probab. 31A (1994), 49–62.
- [8] D. K. Dimitrov, *A refinement of the Gauss-Lucas theorem*, Proc. Amer. Math. Soc., 126(7):2065–2070, 1998.
- [9] C. Döbler, M. Stolz, *Stein’s method and the multivariate CLT for traces of powers on the classical compact groups*, Electronic Journal of Probability 16 (2011), 2375–2405.
- [10] J. Dronka, *On the zeros of a polynomial and its derivative*, Zeszyty Nauk. Politech. Rzeszowskiej. Mat. Fiz. n. 9 (1989), 33–36.
- [11] R. M. Dudley, *Real analysis and probability*, Cambridge, UK, Cambridge University Press (2002).
- [12] G. B. Folland, *A course in abstract harmonic analysis*, Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995.
- [13] J. Fulman, *Stein’s method, heat kernel, and traces of powers of elements of compact Lie groups*, Electronic Journal of Probability, Vol. 17, No. 66, 1–16 (2012).
- [14] A. W. Goodman, Q. I. Rahman, J. S. Ratti, *On the zeros of a polynomial and its derivative*, Proc. Amer. Math. Soc., 21:273–274, 1969.
- [15] J. B. Hough, M. Krishnapur, Y. Peres, B. Virag, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, Volume 51 of University Lecture Series, AMS, Providence, RI, 2009.
- [16] C. P. Hughes, Z. Rudnick, *Mock-Gaussian behaviour for linear statistics of classical compact groups*, Random matrix theory. J. Phys. A 36 (2003), no. 12, 2919–2932.
- [17] K. Johansson, *On random matrices from the compact classical groups*, Ann. of Math. (2) 145 (1997), no. 3, 519–545.

- [18] A. Joyal, *On the zeros of a polynomial and its derivative*, J. Math. Anal. Appl., 26:315–317, 1969.
- [19] Z. Kabluchko, *Critical points of random polynomials with independent identically distributed roots*, Proc. Amer. Math. Soc. **143** (2015), 695–702.
- [20] Z. Kabluchko, D. Zaporozhets, *Asymptotic distribution of complex zeros of random analytic functions*, Ann. Probab. Volume 42, Number 4 (2014), 1374–1395.
- [21] A. Klenke, *Probability Theory*, Springer-Verlag, London, 2014.
- [22] K. Mahler, *On the zeros of the derivative of a polynomial*, Proc. Roy. Soc. Ser. A, 264:145–154, 1961.
- [23] S. M. Malamud, *Inverse spectral problem for normal matrices and the Gauss-Lucas theorem*, Trans. Amer. Math. Soc., 357(10):4043–4064 (electronic), 2005.
- [24] M. Marden, *Geometry of Polynomials*, volume 3 of Mathematical Surveys and Monographs, AMS, 1966.
- [25] M. Marden, *Conjectures on the Critical Points of a Polynomial*, Amer. Math. Monthly, 90(4):267–276, 1983.
- [26] A. Markushevich, *Theory of functions of a complex variable*, Chelsea Publishing Co., New York, 1977.
- [27] E. S. Meckes, M. W. Meckes, *Concentration and convergence rates for spectral measures of random matrices* Probab. Theory Related Fields 156 (2013), no. 1-2, 145–164.
- [28] L. Pastur, V. Vasilchuk, *On the moments of traces of matrices of classical groups*, Comm. Math. Phys. 252 (2004), no. 1-3, 149–166.
- [29] P. Pawłowski, *On the zeros of a polynomial and its derivatives*, Trans. Amer. Math. Soc., 350(11):4461–4472, 1998.
- [30] R. Pemantle, I. Rivin, *The distribution of zeros of the derivative of a random polynomial*, Advances in Combinatorics. Waterloo Workshop in Computer Algebra 2011, I. Kotsireas and E. V. Zima, editors, Springer, New York, 2013.
- [31] V. Petrov, *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*, Oxford Studies in Probability, New York, 1995.
- [32] Q. I. Rahman, *On the zeros of a polynomial and its derivative*, Pacific J. Math., 41:525–528, 1972.
- [33] W. Rudin, *Real and complex analysis (3rd ed.)*, New York: McGraw-Hill (1987).
- [34] B. Sendov, *Hausdorff geometry of polynomials*, East J. Approx., 7(2):123–178, 2001.
- [35] B. Sendov, *New conjectures in the Hausdorff geometry of polynomials*, East J. Approx., 16(2):179–192, 2010.
- [36] A. Soshnikov, *The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities*, Ann. Probab., Volume 28, Number 3 (2000), 1353–1370.
- [37] C. Stein, *The accuracy of the normal approximation to the distribution of the traces of powers of random orthogonal matrices*, Department of Statistics, Stanford University, Technical Report No. 470, 1995.
- [38] M. Stolz, *Stein's method and central limit theorems for Haar distributed orthogonal matrices: some recent developments*, to appear in G. Alsmeyer, M. Löwe (eds.), Random Matrices and Iterated Random Functions, Springer, Proceedings in Mathematics and Statistics.
- [39] È. A. Storozhenko, *A problem of Mahler on the zeros of a polynomial and its derivative*, Mat. Sb., 1996, Volume 187, Number 5, Pages 111–120.
- [40] S. D. Subramanian, *On the distribution of critical points of a polynomial*, Electronic Communications in Probability, Vol 17, No. 37 (2012).
- [41] T. Tao, V. Vu, *Random matrices: Universality of ESDs and the circular law*, Ann. Probab. Volume 38, Number 5 2023–2065 (2010), with an appendix by M. Krishnapur.
- [42] Q. M. Tariq, *On the zeros of a polynomial and its derivative. II*, J. Univ. Kuwait Sci., 13(2):151–156, 1986.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO AT BOULDER, BOULDER, CO 80309
 E-mail address: `sean.d.orourke@colorado.edu`